Exploring How Long Prime Ministers of Australia Lived*

Amie Liu

February 21, 2024

Key Findings

Analyzing the lifespan of Prime Ministers of Australia reveals interesting insights into the longevity and trends among the country's leaders. The dataset contains details regarding the birth years of 31 Prime Ministers, along with their death year and age at death if applicable. Figure 1 shows that out of the 31 Prime Ministers, 8 are currently alive. It covers a broad timeframe, spanning from the late 19th century to the present day (see Figure 2).

The average lifespan of Australian Prime Ministers appears to be relatively high, with many living well into their 70s, 80s, and even 90s. As depicted in Figure 1, Sir Gough Whitlam stands out as the longest-lived Prime Minister, passing away at the remarkable age of 98. Similarly, Sir Billy Hughes, Sir Frank Forde, and Sir John Gorton also lived into their 90s, highlighting exceptional longevity among Australia's leaders. On the other hand, some Prime Ministers had relatively shorter lifespans. Harold Holt, who served as Prime Minister from 1966 and 1967 (Wikipedia 2024), passed away at the age of 59, making him one of the youngest Prime Ministers to die (see Figure 1). Similarly, Joseph Lyons and John Curtin also died relatively young at the age of 60.

Figure 2 presents a clear pattern, illustrating the chronological distribution of prime ministers over time and emphasizing the changes in leadership throughout Australia's history. One notable observation is the steady increase in the number of Prime Ministers over the years, reflecting the growth and evolution of Australia's political landscape. It also highlights periods of stability and continuity in leadership, as well as moments of transition and change. Furthermore, the color-coded differentiation between Prime Ministers who are alive and those

*Code and data are available at: https://github.com/AmieLiu/Exploring-How-Long-Prime-Ministers-of-Australia-lived.git

Prime Minister	Birth year	Death year	Age at death
Edmund Barton	1849	1920	71
Alfred Deakin	1856	1919	63
Chris Watson	1867	1941	74
George Reid	1845	1918	73
Andrew Fisher	1862	1928	66
Joseph Cook	1860	1947	87
Billy Hughes	1862	1952	90
Stanley Bruce	1883	1967	84
James Scullin	1876	1953	77
Joseph Lyons	1879	1939	60
Earle Page	1880	1961	81
Robert Menzies	1894	1978	84
Arthur Fadden	1894	1973	79
John Curtin	1885	1945	60
Frank Forde	1890	1983	93
Ben Chifley	1885	1951	66
Harold Holt	1908	1967	59
John McEwen	1900	1980	80
John Gorton	1911	2002	91
William McMahon	1908	1988	80
Gough Whitlam	1916	2014	98
Malcolm Fraser	1930	2015	85
Bob Hawke	1929	2019	90
Paul Keating	1944	NA	NA
John Howard	1936	NA	NA
Kevin Rudd	1957	NA	NA
Julia Gillard	1961	NA	NA
Tony Abbott	1957	NA	NA
Malcolm Turnbull	1954	NA	NA
Scott Morrison	1968	NA	NA
Anthony Albanese	1963	NA	NA

Figure 1: Table of Australian Prime Ministers

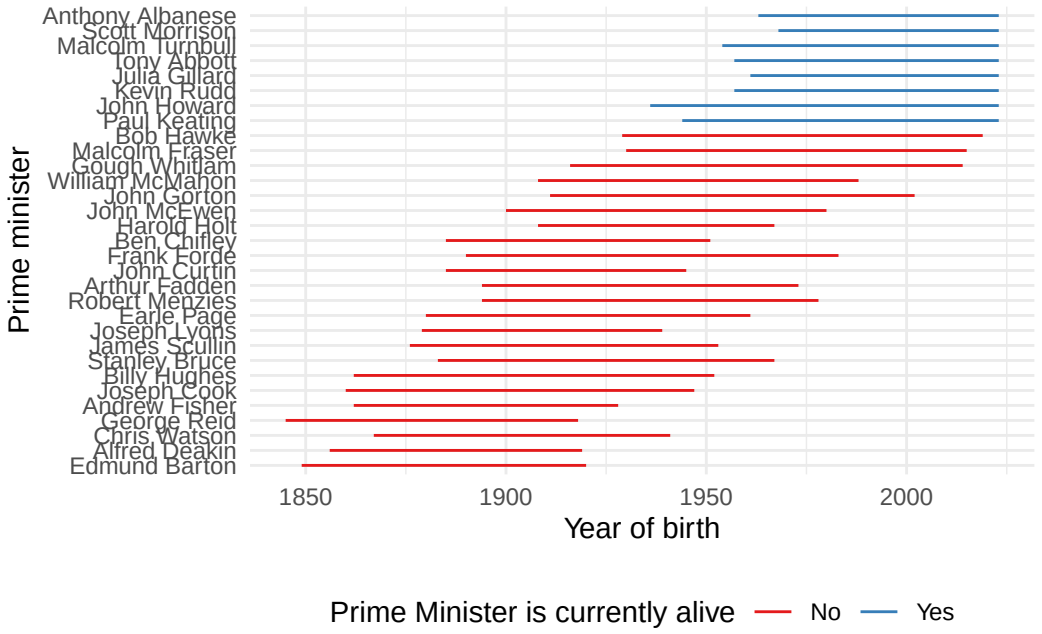

Figure 2: Lifespan of Australian Prime Ministers

who have passed away provides insights into their current status (see Figure 2). This information enhances our understanding of Australia's political history and the legacies of its past leaders.

A Data Journey from Wikipedia to Visualization with R

The data on Australian Prime Ministers was gathered through web scraping from Wikipedia (Wikipedia 2024) and was then created into a table and graph, using the programming language R (R Core Team 2022), 'tidyverse' (Wickham et al. 2019), 'janitor' (Firke 2023), 'ggplot2' (Wickham 2016), 'babynames' (Wickham 2021), 'rvest' (Wickham 2024), and 'xml2' (Wickham, Hester, and Ooms 2023) packages. This source provided details about each prime minister, including their names, birth years, death years (if applicable), and other relevant biographical information.

Web scraping using the 'rvest' package (Wickham 2024) allows for the automated extraction of structured data from web pages, making it an efficient and effective method for gathering large datasets from online sources. By utilizing web scraping techniques with the 'rvest' package (Wickham 2024), I was able to access and gather data on Australian Prime Ministers directly from Wikipedia, ensuring the accuracy and reliability of the information obtained.

After gathering the data, I followed the procedures from "Telling Stories with Data: 7 Gather Data" (Alexander, n.d.) to simulate the dataset, clean the data, and create a table and graph. I first simulated the data with four columns of variables, including Prime minister, Birth year, Death year, and Years lived. Then, I standardized naming conventions to ensure consistency in formatting, handled missing values, and removed irrelevant information. Once the data was cleaned, I created a table and graph similar to the simulated data, with columns for prime minister names, birth years, death years, and age at death.

Reflection

One challenge faced during the process was generating the graph, especially when some birth years were missing after data cleaning. In this case, I had to carefully search through Wikipedia (Wikipedia 2024) to fill in the missing information. This process was time-consuming to ensure the dataset was as comprehensive as possible.

Despite the challenges, the process became increasingly enjoyable as I delved deeper into the historical records and biographical details of Australian prime ministers. Discovering insights about their lives and contributions to the national political landscape was both informative and engaging, making the task of gathering and cleaning data an enjoyable experience.

Reflecting on this paper, there are several aspects I would approach differently in future data gathering and cleaning. Firstly, I would allocate more time for data verification to ensure the highest level of accuracy and reliability. Additionally, I would utilize more advanced data cleaning techniques to automate certain aspects of the cleaning process, streamline data preparation, and increase data visualization. Furthermore, I would prioritize creating detailed documentation throughout the data gathering and cleaning process, documenting sources, methodologies, and any challenges or insights encountered along the way. This would facilitate transparency and reproducibility.

References

Alexander, Rohan. n.d. "7 Gather Data." *Telling Stories with Data*. https://tellingstorieswithdata.com/07-gather.html#prime-ministers-of-the-united-kingdom.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

———. 2021. *Babynames: US Baby Names 1880-2017*. https://CRAN.R-project.org/package=babynames.

———. 2024. *Rvest: Easily Harvest (Scrape) Web Pages*. https://CRAN.R-project.org/package=rvest.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Jim Hester, and Jeroen Ooms. 2023. *Xml2: Parse XML*. https://CRAN.R-project.org/package=xml2.

Wikipedia. 2024. *Wikipedia*. Wikimedia Foundation. https://en.wikipedia.org/wiki/List_of_prime_ministers_of_Australia.